Building intelligent, scalable, high performance NVMe based storage systems with smart PCIe 4.0 embedded switch IP

Sampath Banka

Agenda

- Introduction: Storage systems and PCIe
- Vendor differentiation through SoC integration
- Vendor differentiation with PCIe switching
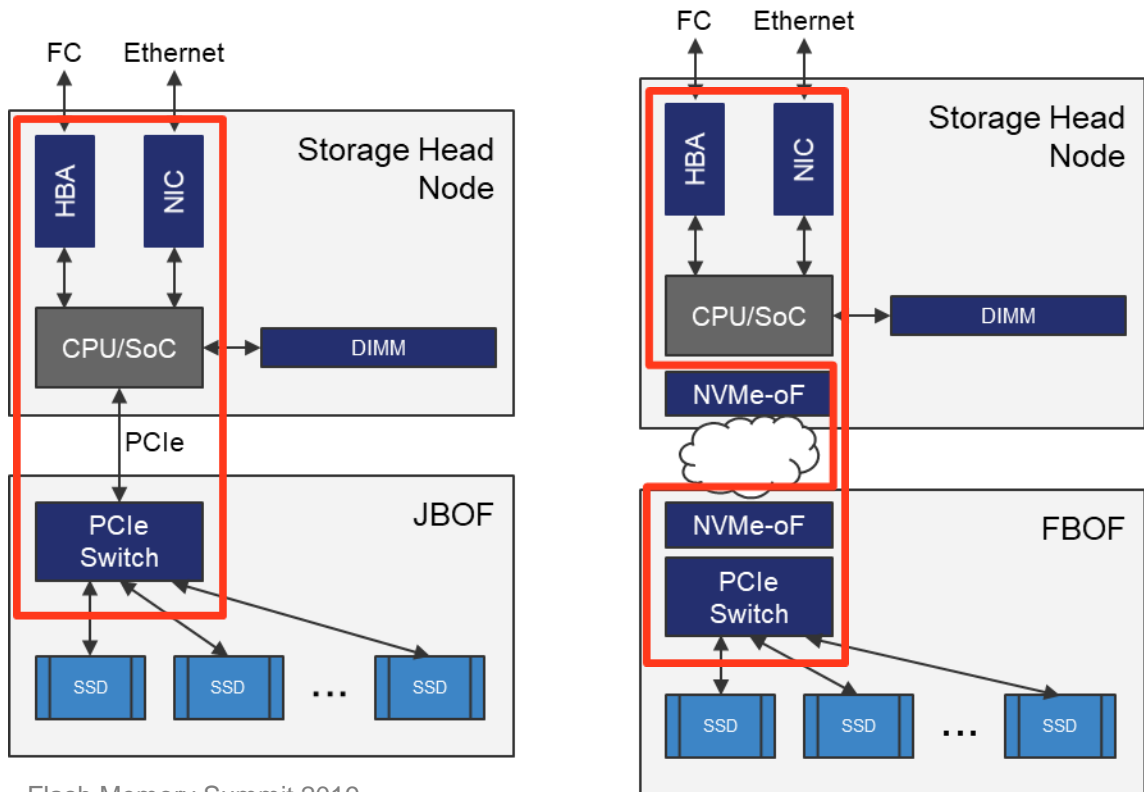- Embedded switch IP architecture and benefits
- Conclusion

Introduction

- Storage systems (direct attached, disaggregated) rely on PCIe switching for NVMe SSD fanout
- PCIe switching based on commodity switch chips
  - One size fits all model
  - BoM/cost/power consideration
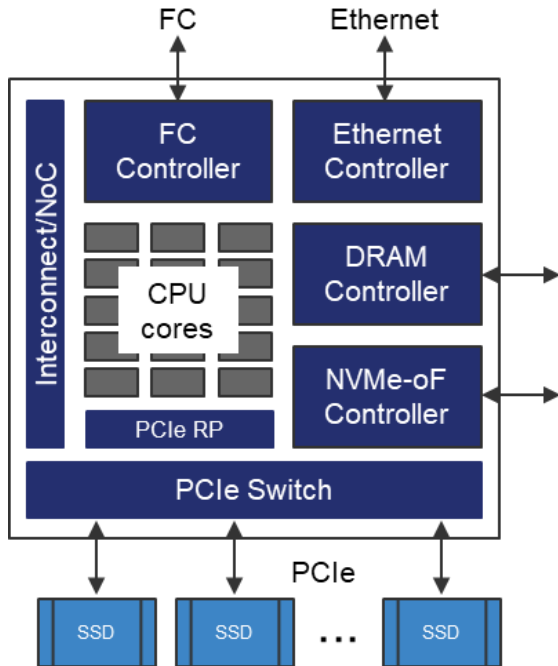 - Little/no differentiation at the HW level

Typical Storage Architectures

- Direct-attached vs. fabric-attached

- Discrete PCIe Switches

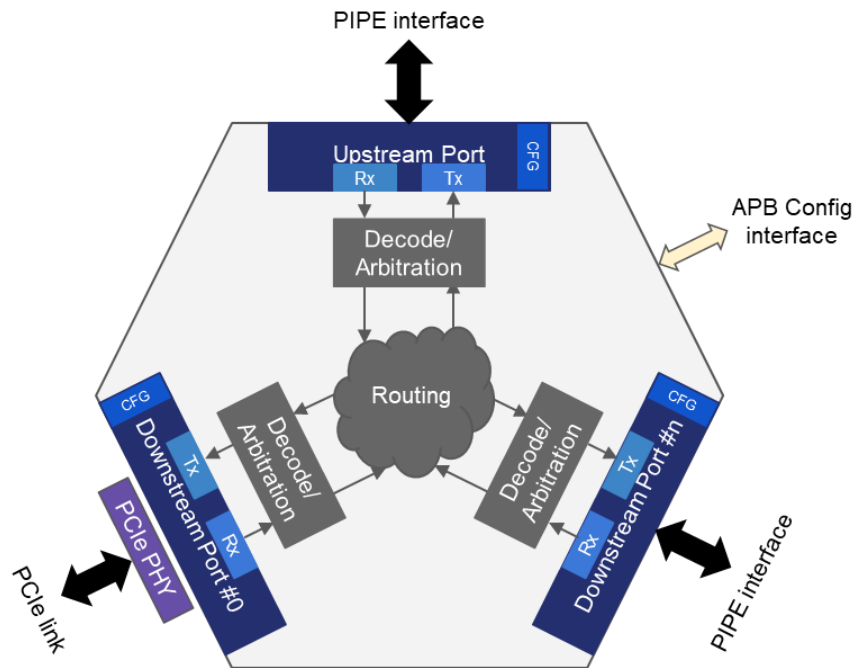- Vendor differentiation at the Software level

Differentiation with Silicon Integration

- Towards one-SoC-fits-all-architectures
  - Can serve as Host controller or Flash controller
  - Improved performance (no discrete components)
  - Optimized BoM, power, features
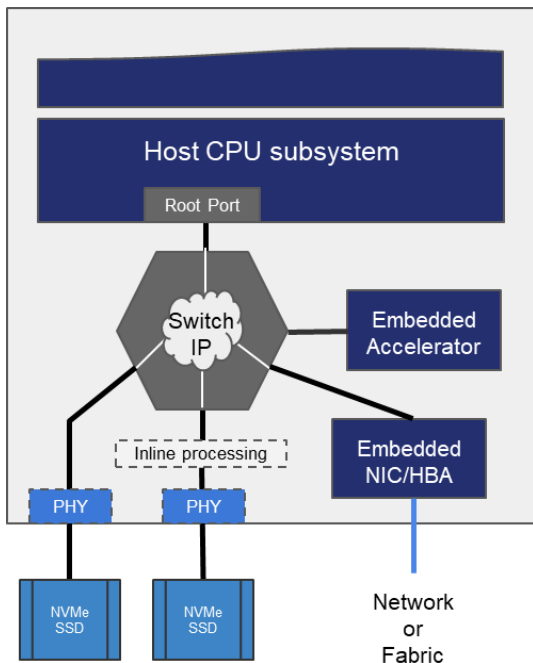 - Future-proof with latest PCIe gen

PCIe Switch IP Architecture

- 1 Upstream (single host), multiple Downstream ports (fanout switch)
- External connections via PCIe PHY
- Internal connections via PIPE
  - Reduced latency, gate count, power
- Transparent design means NVMe SSDs "seen" by Host

Differentiating with PCIe Switch IP

- With embedded endpoints
  - Accelerators for HW-based data processing (encryption, compression)
  - HBAs and NICs for lower latency, lower power (no PHY), reduced BoM
  - Switch needs P2P support
- With inline processing capabilities
  - Ex. filtering, snooping/statistics
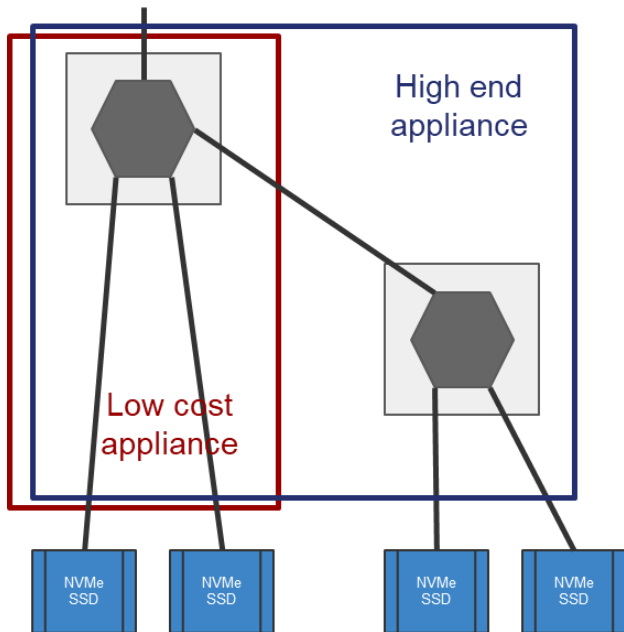 - Switch needs pipelining support

Multi-Host Architectures
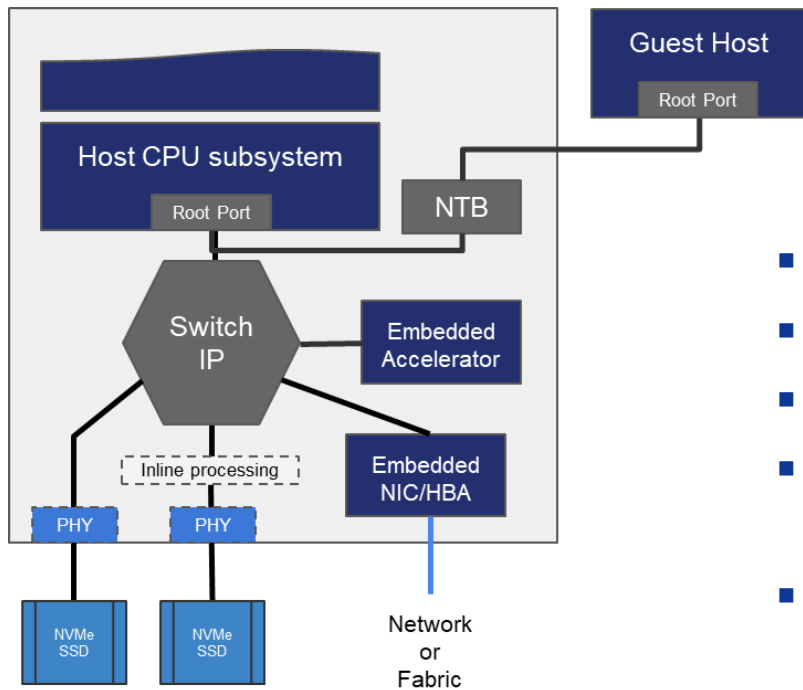
- PCIe Switch IP with NTB capability
- Can add multiple guest Hosts
- Host to Host comm. Via NTB
- Guest Host to Device comm. via Address Translation (ATL)
- NTB requires custom device driver

Switch Cascading

- Allows differentiation at the appliance level
  - Modular design: from low-cost to high-end
  - Multiple instances of same SoC/FPGA

Top eSwitch Features for Storage

- Low latency switch (cut-through architecture)

- Support of PCIe 4.0 or 5.0 and SR-IOV

- RAS features incl. ECRC, Parity, ECC, AER, Hot Plug

- Support for inline and/or embedded processing via pipelining and/or embedded endpoints

- Any port to any port high-speed communication including peer-to-peer between endpoint devices

- Ability to support different link width and speed on downstream ports

- Downstream Port Containment (DPC)

Conclusion

- Storage systems trending towards more SoC integration for differentiation

- PCIe switching is a key component, requires well architected PCIe switch IP

- Embedded switch IP provides HW-level differentiation with embedded endpoint, inline processing

- Embedded switch IP provides system-level benefits such as lower latency, reduced BoM/costs, PCIe future-proofing

Thank you !

Any questions ?